

A unified approach to FFT based image registration

A. Averbuch¹, Y. Keller^{1,2}

¹School of Computer Science, Tel Aviv University, Tel Aviv 69978

Israel

²Dept. of Electrical Engineering Systems

Tel-Aviv University

Tel-Aviv 69978, Israel

Abstract

We present a new unified approach to FFT based image registration. Prior works divided the registration process into two steps: the first was based on phase correlation (PC) which provides pixel accurate registration [7], while the second step provides subpixel registration accuracy [2, 5]. By extending the PC method we derive a FFT based image registration algorithm which is able to estimate large translations with subpixel accuracy. The algorithm's properties resemble those of the Gradient Methods [6] while outperforming them by exhibiting superior convergence range. Furthermore, the algorithm is robust when registering severely noisy images.

Keywords: Sub-pixel registration, phase correlation, image alignment

EDICS Category={2-ANAL, 2-MOTD}

1 Introduction

Image registration is an essential component in many computer vision and image processing tasks such as motion analysis, video compression, image enhancement and restoration. Many of these tasks require pixel accurate registration (motion detection) while others (video compression, image enhancement, stereo vision) require sub-pixel accuracy. Gradient based methods (GM) [6] are considered to be the state-of-the-art registration algorithms, being able to account for both pixel and sub-pixel registration in various motion models (translation, rotation, affine, projective). The major drawback of GM is its need to use bootstrapping due to its limited convergence range, even when implemented in a multi-resolution setup. Thus, large translations have to be initially estimated before invoking the GM. A possible bootstrapping algorithm is the phase correlation (PC) [7] scheme which utilizes the phase shift property [1] of the Fourier transform to robustly estimate large image translations without prior knowledge. It is usually limited to translation while it is possible (through Cartesian \rightarrow Polar interpolation [4]) to estimate scale changes and rotations. An extension to sub-pixel registration was presented by Shekarforoush et-al [2], while it is strictly limited to sub-pixel shifts and has to be bootstrapped using pixel accurate PC. We propose to use frequency analysis to derive a Unified PC (UPC) image registration algorithm being able to estimate large translations with sub-pixel accuracy. We utilize the notion that the phase shifts between continuous signals are of infinite accuracy and are invariant under the sampling process. Further analysis reveals a close relationship to GM methods, while exhibiting superior analytic properties.

The paper is organized as follows: Problem formulation and basic structure of the pixel

accurate PC are given in section 2. Fourier domain analysis leading to a sub-pixel registration algorithm are given in section 3. Implementation issues and results are discussed in 4. The relationship of the methods to GM is given in section 5 and Appendix A.

2 Background

2.1 Pixel Level Phase Correlation

2.1.1 Phase correlation

The purpose of the PC algorithm is to estimate the translation $(\Delta x, \Delta y)$ between a pair of images sharing some common support: Let $I_1(x, y)$ and $I_2(x, y)$ be the images described in Fig. 1, then their common support satisfies:

$$I_1(x + \Delta x, y + \Delta y) = I_2(x, y) \tag{2.1}$$

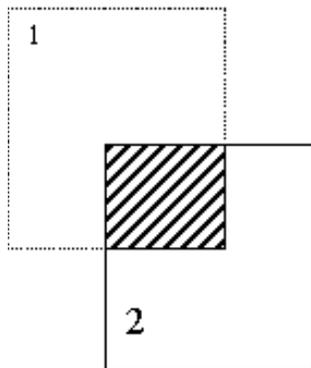


Figure 1: Image translation

Equation 2.1 can be generalized to account for affine intensity changes:

$$I_1(x + \Delta x, y + \Delta y) = a \cdot I_2(x, y) + b, \quad a, b \in \mathfrak{R} \quad (2.2)$$

Denote by $\mathfrak{F}\{f(x, y)\}$ the Fourier transform of $f(x, y)$:

$$\mathfrak{F}\{f(x, y)\} \triangleq \hat{f}(\omega_x, \omega_y) \quad (2.3)$$

Equation 2.2 is solved by the PC algorithm using the *shift property* [1] of the FFT:

Lemma 1 *Let $f(x, y)$ and $\hat{f}(\omega_x, \omega_y)$ be a function on \mathfrak{R}^2 and its 2D Fourier transform, respectively.*

Then

$$\mathfrak{F}\{f(x + \Delta x, y + \Delta y)\} = \hat{f}(\omega_x, \omega_y) e^{j(\omega_x \Delta x + \omega_y \Delta y)} \quad (2.4)$$

We transform Eq. 2.1 into the Fourier domain using Eq. 2.4:

$$\hat{I}_1(\omega_x, \omega_y) e^{j(\omega_x \Delta x + \omega_y \Delta y)} = \hat{I}_2(\omega_x, \omega_y) \quad (2.5)$$

$$\frac{\hat{I}_2(\omega_x, \omega_y)}{\hat{I}_1(\omega_x, \omega_y)} = e^{j(\omega_x \Delta x + \omega_y \Delta y)} \quad (2.6)$$

The translation $(\Delta x, \Delta y)$ can be estimated by taking the inverse FFT of Eq. 2.6 :

$$Corr(x, y) \triangleq \mathfrak{F}^{-1}\{e^{j(\omega_x \Delta x + \omega_y \Delta y)}\} = \delta(x - \Delta x, y - \Delta y) \quad (2.7)$$

Ideally, the peak value of the correlation surface $Corr(x, y)$ should be equal to 1.0. However, in real scenarios, due to the presence of random noise, dissimilar parts and non-translational motion the peak value will be smaller than 1.0. Therefore, the translation values are estimated by:

$$(x, y) = \arg \left\{ \max_{(\tilde{x}, \tilde{y})} \{Corr(\tilde{x}, \tilde{y})\} \right\}. \quad (2.8)$$

In order to compensate for possible gain changes (Eq. 2.2) and since we are only interested in pure phase shifts, Eq. 2.5 can be rewritten as:

$$\widehat{Corr}(\omega_x, \omega_y) = \frac{\widehat{f}_1(\omega_x, \omega_y) \widehat{f}_2^*(\omega_x, \omega_y)}{\left| \widehat{f}_1(\omega_x, \omega_y) \right| \left| \widehat{f}_2(\omega_x, \omega_y) \right|} = e^{j(\omega_x \Delta x + \omega_y \Delta y)} \quad (2.9)$$

where * denotes the complex conjugate.

Equation 2.9 is an estimator of the correlation surface which can be interpreted as a projection operator, which projects the result of Eq. 2.6 onto the space of phase shift functions, since it guaranties that $\left| \widehat{Corr}(\omega_x, \omega_y) \right| = 1$.

2.2 Sub-pixel registration

Shekarforoush et-al [2] presented a state-of-the-art interpolation-free, PC based, sub-pixel registration algorithm by analyzing the effects of sub-pixel shifts on the correlation function $Corr(x, y)$ in the *spatial* domain. Their approach is briefly presented in section 2.2.1. In section 3 we propose to analyze $Corr(x, y)$ in the *frequency* domain, thus achieving much superior sub-pixel accuracy and convergence range.

2.2.1 Spatial domain sub-pixel registration

In Shekarforoush et-al [2], bootstrapping is achieved via the calculation of the pixel accurate PC (section 2.1), thus restricting the registration uncertainty to sub-pixel shifts of magnitudes $(|\Delta x|, |\Delta y|) < 1$. Refinement to sub-pixel accuracy is achieved by analyzing the 1D inverse Fourier transform of a single frequency exponent, whose phase-shift relates to sub-pixel spatial shifts:

$$\hat{f}(i) = \{e^{ji\Delta\omega\Delta t}\}_i, \quad 0 \leq i < m, \quad \Delta\omega = \frac{2\pi}{m}, \quad \Delta t \notin \mathbb{Z} \quad (2.10)$$

The result $f(k)$ is the Dirichlet Kernel $D(\Delta t + k, m)$ (illustrated by Fig. 2):

$$D(\Delta t + k, m) = \frac{\sin(\pi(\Delta t + k))}{\sin(\pi(\Delta t + k)/m)} \quad (2.11)$$

$D(\Delta t + k, N)$ coincides with $\delta(x, y)$ (in Eq. 2.7) for integer pixel shifts while for sub-pixel shifts the center lobe's energy is split into side-lobes.

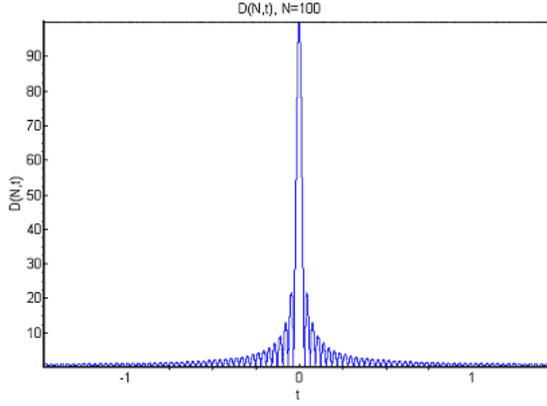


Figure 2: Derichle kernel

Using the separability of the 2D DFT this analysis can be conducted on both axes, yielding:

$$\frac{\sin(\pi(\Delta x + x))}{\sin(\pi(\Delta x + x)/m)} \cdot \frac{\sin(\pi(\Delta y + y))}{\sin(\pi(\Delta y + y)/n)} = D(x, y, m), (x, y) \in \mathbb{Z} \quad (2.12)$$

where m, n are the image dimensions and $\Delta x, \Delta y$ are the sought after translation values.

The non-linear Eq. 2.12 is solved for $(\Delta x, \Delta y)$ using linearization of the denominator around $(x, y) = (0, 0)$: assuming $(m, n) \gg 1$ (since $(|\Delta x|, |\Delta y|) < 1$) yields $\pi(\Delta y + 0)/n \ll 1, \pi(\Delta x + 0)/m \ll 1$, therefore we have

$$\begin{aligned} \sin(\pi\Delta y/n) &\approx \pi\Delta y/n \\ \sin(\pi\Delta x/m) &\approx \pi\Delta x/m. \end{aligned} \quad (2.13)$$

By substituting Eq. 2.13 into 2.12 we get a simplified linearized equation:

$$\frac{\sin(\pi\Delta x)}{\pi\Delta x/m} \cdot \frac{\sin(\pi\Delta y)}{\pi\Delta y/n} = D(x, y, m), (x, y) \in \mathbb{Z}. \quad (2.14)$$

An estimate of $(\Delta x, \Delta y)$ is derived by substituting the correlation surface's main peak and neighboring peaks into Eq. 2.14:

$$\begin{cases} \Delta x = \frac{C(1,0)}{C(1,0) \pm C(0,0)} \\ \Delta y = \frac{C(0,1)}{C(0,1) \pm C(0,0)} \end{cases} \quad (2.15)$$

where $C(0,0)$ is the correlation surface's main peak located at $(0,0)$ and $C(1,0)$ and $C(0,1)$ are the neighboring vertical and horizontal side-lobes, respectively. According to Eq. 2.15 the accuracy of the estimate is linearly dependent upon the measurement accuracy of the side lobes, thus limiting the estimation accuracy, since the ratio of the highest side lobe to the main lobe is 21% (-13.5dB) [1]. Performance comparison between the results of [2] and the proposed UPC is presented in section 4.3.

3 Sub-pixel registration in the frequency domain

The motivation that led to the proposed UPC sub-pixel registration is the notion that Eq. 2.5 holds for any non-integer displacement values and its accuracy is only limited by errors induced by the problem setup: finite image size, non-overlapping areas and illumination changes between the images which are not modeled by Eq. 2.2. By calculating the phase shift directly, one can estimate an algebraically correct phase shift which is not limited to integer offsets or pixel accuracy.

3.1 Flow of the proposed algorithm

The proposed sub-pixel registration computation process goes as follows:

Step 1: Calculate the correlation surface $Corr(x, y)$ using Eq. 2.9 and project it into the appropriate registration uncertainty domain (section 3.2).

Step 2: Calculate the pointwise phase shift functions in the Fourier domain using the Fourier coefficients (section 3.3).

Step 3: Using the robust estimator described in section 3.4, estimate the global sub-pixel accurate translation $(\Delta x, \Delta y)$.

Step 4: Estimate and extract the common areas \tilde{I}_1 and \tilde{I}_2 is created by the pixel level motion $(\lfloor \Delta x \rfloor, \lfloor \Delta y \rfloor)$ according to Fig. 1.

Step 5: Move \tilde{I}_2 by the sub-pixel motion $(\Delta x - \lfloor \Delta x \rfloor, \Delta y - \lfloor \Delta y \rfloor)$ (section 3.5).

Step 6: Go back to step # 2 until a stopping condition is met. Our experience shows that a few iterations (usually 5-8 iterations) are sufficient for convergence to $O(10^{-3})$ accuracy.

3.2 Phase projection operator

Following the results of section 2.2.1, the correlation surface $Corr(x, y)$ is expected to have a limited spatial support of 5–7 pixels. Non-zeros elements beyond this support are a result of noisy measurements and should be filtered out. Hence, the phase-shift estimation can be improved by limiting the correlation surface’s support to a rectangular domain of dimensions

$[-R, R] \times [-R, R]$. Next we provide an analytic analysis of the estimation improvement, while experimental results are provided in section 4.4.

3.2.1 Phase projection operator gain analysis

Denote by $\widetilde{Corr}(\omega_x, \omega_y)$ the measured correlation surface. Then, by using Eq. 2.5 we have:

$$\widetilde{Corr}(\omega_x, \omega_y) = \widehat{Corr}(\omega_x, \omega_y) + \widehat{n}(\omega_x, \omega_y) \quad (3.1)$$

where $\widehat{Corr}(\omega_x, \omega_y)$ is the true correlation surface and $\widehat{n}(\omega_x, \omega_y)$ is WGN (white gaussian noise). $\widehat{n}(\omega_x, \omega_y)$ is distributed evenly in the frequency and spatial domain, thus, by limiting the support of $\widetilde{Corr}(x, y)$ to $[-R, R] \times [-R, R]$ in the *spatial domain* the WGN energy is decreased by a factor of $\frac{N \cdot M}{R^2}$ where N and M are the dimensions of the $\widetilde{Corr}(\omega_x, \omega_y)$. Approximately 90% of the energy of $Corr(x, y)$ (section 2.2.1) is concentrated in the domain $[-R, R] \times [-R, R]$, hence the SNR gain is:

$$\Delta SNR \approx \frac{N \cdot M}{R^2} \times 0.9 \quad (3.2)$$

For a typical image of size 512×512 and $R = 5$ (a typical effective size used in our experiments) this results in a significant SNR improvement:

$$\Delta SNR \approx \frac{512 \cdot 512}{25} \times 0.9 = 9,437.2 = 39.74dB \quad (3.3)$$

3.3 Phase estimation

The Fourier based correlation $\widehat{Corr}(\omega_x, \omega_y)$ is calculated using Eq. 2.5, while the pointwise phase is estimated using the following computation:

$$\varphi(i, j) = \tan^{-1} \left\{ \frac{\text{Im} \left(\widehat{Corr}(\omega_x, \omega_y) \right)}{\text{Re} \left(\widehat{Corr}(\omega_x, \omega_y) \right)} \right\} \quad (3.4)$$

where:

$$\begin{cases} \omega_x = j \cdot 2\pi/m \\ \omega_y = i \cdot 2\pi/n \end{cases} \quad i, j \in \text{support} \left(\widehat{Corr}(\omega_x, \omega_y) \right). \quad (3.5)$$

This phase estimation is invariant under affine illumination changes since the phase is solely used. Gain differences do not affect the phase shift due to the normalization in Eq. 2.5, while offset changes affect only the DC magnitude but not its phase. Furthermore, this method is immune to image blurring of one of the images by linear filters (which do not change the phase) or both images by the same linear filter. In this case the phase of both images is equally changed.

3.4 Robust least squares estimator

Using the $\varphi(i, j)$ estimation in Eq. 3.4 we derive a highly overdetermined set of linear equations

$$\varphi(i, j) = \omega_x \Delta x + \omega_y \Delta y \quad i, j = 1 \dots n, m \quad (3.6)$$

where $(\Delta x, \Delta y)$ are the sought after translations.

Equation 3.6 can be solved directly using least square (LS) which was found to perform well in noise-free scenarios (section 4.4). However, in order to register noisy images with significant dissimilarities, a robust LS estimator has to be introduced. The robust estimator aims to classify each phase shift measurement as belonging to either “inliers” or “outliers”, where the outliers will not be included in the shift regression. Following [3] an iteratively re-weighted, least squares regression algorithm is used: at each iteration a weighting function is applied to the residuals of the previous iteration, assigning lower weights to points that do not fit well. Next, the weights are used to re-calculate a weighted LS estimate. Several weight functions were tested yielding very similar results.

3.5 Induced Sub-pixel Translation

In most iterative sub-pixel registration schemes including the GM, sub-pixel translation is performed using bilinear interpolation. We found it to be inadequate due to the non-linear shift induced at high-frequencies. An example is given in Fig. 3 which depicts the phase response (upper figure) and phase shift (lower figure) of a mask corresponding to bilinear shift of $(\Delta x, \Delta y) = (0.5, 0.5)$. The phase is linear for the low frequencies, while a phase reversal occurs in the high frequencies due to the existence of a zero in the transfer function of the bilinear interpolation.

In order to implement a sub-pixel translation we utilize a Fourier domain phase shift based on Eq. 2.5, similar to the one used in [8]: each 2D Fourier coefficient is multiplied by an appropriate complex phase shift coefficient calculated numerically:

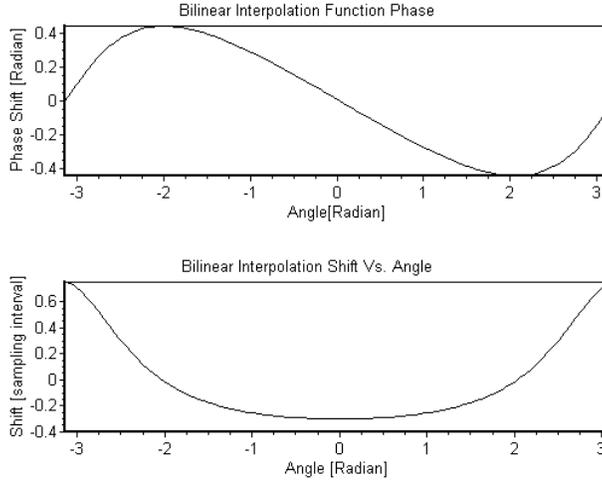


Figure 3: The phase shift after bilinear interpolation: the upper picture shows the phase of bilinear interpolation which is non-linear, while the lower picture shows the induced phase-shift which is non-constant.

$$\widehat{I}_{shift}(\omega_x, \omega_y) = \widehat{I}_{base}(\omega_x, \omega_y) e^{j(\omega_x \widetilde{\Delta x} + \omega_y \widetilde{\Delta y})} \quad (3.7)$$

where:

\widehat{I}_{base} is the source image

\widehat{I}_{shift} is the resulted shifted image

$(\widetilde{\Delta x}, \widetilde{\Delta y})$ are the induced shifts.

Although Eq. 3.7 can theoretically be used to cause shifts of any magnitude, integer shifts ($\lfloor \Delta x \rfloor, \lfloor \Delta y \rfloor$) were implemented using common support extraction according to Fig. 1, thus achieving ultimate accuracy through a simple implementation. The residual shifts $(\Delta x - \lfloor \Delta x \rfloor, \Delta y - \lfloor \Delta y \rfloor)$ were induced using Eq. 3.7.

4 Experimental Results

In order to demonstrate the performance of the algorithm, the images shown in Fig. 4 were used: Fig. 4(a),(b) and (c) are airborne images of various types. (a) is a textured image lacking distinct features while (b) and (c) contain dominant man-made objects. Image (d) (“Barbara”) is a typical multimedia image.

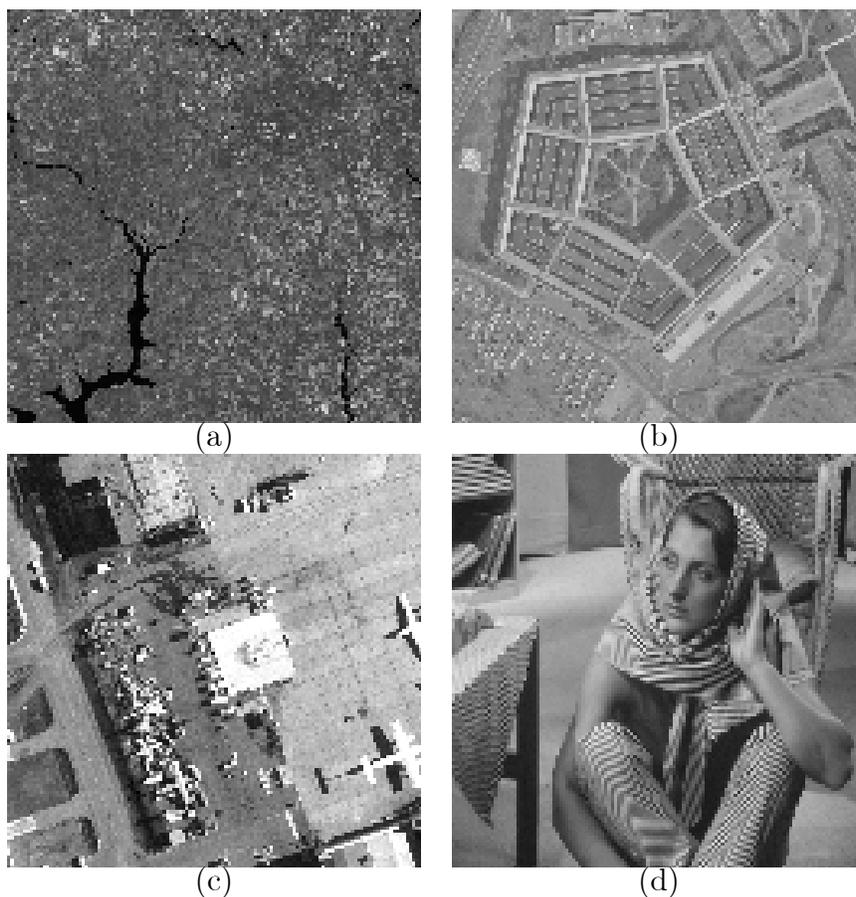


Figure 4: Test images used: (a) SAR image, (b) “Pentagon”, (c) ”Airfield” and (d) ”Barabara”

4.1 Generation of sub-pixel shifted test images

The images were shifted using a procedure similar to the one used in [2]. Let:

$I_1(x, y)$ and $I_2(x, y)$ be the input images.

$(\Delta x, \Delta y) = \left(\frac{k_1}{m_1}, \frac{k_2}{m_2}\right)$ are the required sub-pixel shifts, where $(k_1, k_2, m_1, m_2) \in \mathbb{Z}$, then:

Step 1: $I_1(x, y)$ and $I_2(x, y)$ are smoothed using a Gaussian kernel to avoid aliasing caused by the downscaling in later steps..

Step 2: Using Fig. 1, extract $\tilde{I}_{base}(x, y)$ and $\tilde{I}_{shift}(x, y)$, the common supports of $I_1(x, y)$ and $I_2(x + k_1, y + k_2)$, respectively.

Step 3: $\tilde{I}_{base}(x, y)$ and $\tilde{I}_{shift}(x, y)$ are downsampled by the factors (m_1, m_2) in (x, y) , respectively, to generate the shifted images.

An example of sub-pixel shifted images is given in Fig. 5.



Figure 5: Sub-pixl shifted test images: left image is shifted by $(\Delta x, \Delta y) = \left(\frac{1}{3}, -\frac{2}{3}\right)$ relatively to the right image.

4.2 Registration accuracy estimate

Estimation of large shifts by the proposed algorithm was tested using the images in Fig. 4.

They were shifted by as much as 20 pixels having a sub-pixel component. The results, which

are displayed in Table 1, show that the average UPC registration error was of $O(10^{-3})$ pixel and convergence was achieved in 6–7 iterations. The robust LS, compared well to the regular LS, showing superior accuracy of $O\left(\frac{5}{1000}\right)$ pixel at a cost of a significant complexity increase. Hence, it seems that the regular LS is better suited for most practical scenarios. However, our results suggest that when a significant amount of noise or dissimilarity between the input images exists, the robust LS should be considered (see section 4.4).

Image	The true ($\Delta x, \Delta y$)	UPC using robust LS	Error [10^{-3}]	n	UPC using regular L	Error [10^{-3}]	n
Pentagon	(2.50, -2.50)	(2.500, -2.500)	(0, 0)	5	(2.501, -2.501)	(1, 1)	6
	(3.25, 1.50)	(3.249, 1.505)	(1, 5)	6	(3.248, 1.513)	(2, 13)	8
	(-5.25, -4.50)	(-5.249, -4.509)	(1, 9)	8	(-5.248, -4.510)	(2, 10)	8
	(0.0, 15.75)	(-0.000, 15.752)	(0, 3)	8	(-0.000, 15.752)	(0, 2)	6
Airfield	(0.167, -0.50)	(0.170, -0.504)	(3, 4)	5	(0.170, -0.504)	(3, 4)	5
	(0.5, 3.25)	(0.500, 3.252)	(0, 2)	5	(0.501, 3.251)	(1, 1)	5
	(-7.33, -1.167)	(-7.338, -1.169)	(5, 2)	8	(-7.334, -1.168)	(1, 1)	6
	(-7.33, 2.83)	(-7.337, 2.838)	(7, 5)	8	(-7.336, 2.839)	(3, 6)	6
Sar	(-7.33, 2.83)	(-7.337, 2.838)	(4, 5)	7	(-7.33, 2.835)	(1, 2)	7
	(6.33, 7.33)	(6.334, 7.334)	(1, 1)	8	(6.34, 7.335)	(7, 2)	6
	(-7.66, 21.66)	(-17.669, 21.671)	(3, 5)	8	(-17.669, 21.672)	(3, 6)	7
	(-17.66, -0.33)	(-17.6708, -0.3362)	(10, 3)	8	(-17.670, -0.337)	(1, 4)	5
Barbara	(-7.33, 2.83)	(-7.336, 2.840)	(3, 7)	7	(-7.335, 2.839)	(2, 6)	7
	(6.33, 7.33)	(6.338, 7.340)	(5, 7)	9	(6.340, 7.3419)	(1, 8)	7
	(-17.66, 21.66)	(-17.674, 21.679)	(8, 13)	9	(-17.673, 21.685)	(7, 19)	9
	(-17.66, -0.33)	(-17.673, -0.338)	(7, 5)	7	(-17.674, -0.340)	(8, 10)	5

Table 1: Estimation results for large shift. n is the number of iterations used.

4.3 Registration accuracy comparison

We compare our results to those reported by Shekarforoush et-al [2] by recreating the shifted “Pentagon” image used in it. The results, summarized in Table 2, point out that our algorithm significantly outperforms it by using either the robust or the regular LS. Note that the algorithm in [2] is unable to directly estimate shifts larger than one pixel. Thus, no comparisons with large shifts were needed. The use of the robust LS results in a higher complexity

of the UPC compared to [2]. However, when the regular LS is used, the complexities of the UPC and [2] algorithms are dominated by the FFT algorithm and are approximately the same.

Image	$(\Delta x, \Delta y)$	Shekarforoush [2]	Error [10^{-4}]	Unified PC 3 iterations	Error [10^{-4}]
Pentagon	(0.50, -0.50)	(0.48, -0.51)	(200, 100)	(0.4997, -0.5005)	(3, 5)
	(0.25, 0.50)	(0.28, 0.49)	(300, 100)	(0.2499, 0.5008)	(1, 8)
	(-0.25, -0.50)	(-0.25, -0.52)	(0, 200)	(-0.2504, -0.5002)	(4, 2)
	(0.0, 0.75)	(0.0, 0.80)	(200, 500)	(-0.0006, 0.7510)	(6, 10)

Table 2: Comparison between the UPC’s sub-pixel shift estimation and the results reported by Shekarforoush[2]. We were unable to retrieve the other test images in [2].

4.4 Registration robustness test

The choice of the phase projection operator (section 3.2) was validated by experiments carried out on simulated examples where WGN was added to the shifted image. They revealed that for noise-free images, such as the ones used in section 4.2, there was no need for the operator since the robust LS and regular LS were able to handle the associated problem with setup noise. However, when a significant amount of noise was added, as in Fig. 6 (WGN, $\sigma = 20$), the radius R of the projection operator dominated the results presented in Table 3. For $R < 5$ the UPC failed to converge, while for $5 \leq R \leq 21$ a registration accuracy of $O(10^{-2})$ was achieved without lowering the convergence range. When the projection operator was not used, the robust LS yielded errors of $O(10^{-1})$ while the regular LS failed to converge.

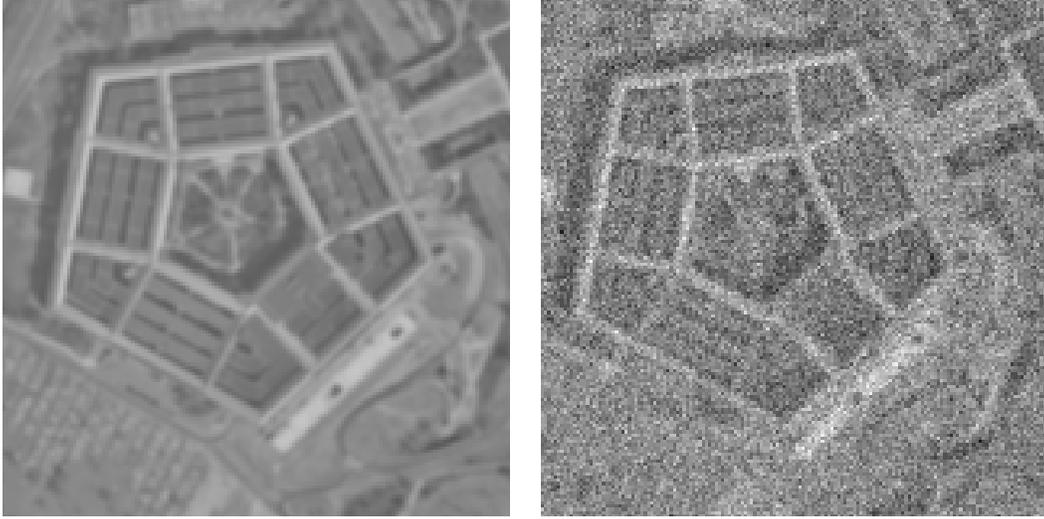


Figure 6: Robustness test images: (a) original “Pentagon” image, (b) “Pentagon” image shifted by $(\Delta x = 5.0, \Delta y = -2.33)$ with added WGN noise $(\sigma = 20, \mu = 0)$

Image	Robust regression	Correlation surface extraction radius	Estimated $(\Delta x, \Delta y)$	Iteration#
Pentagon	✓	3	(0.12, 2.68)	15
	✓	5	(5.01, -2.31)	8
	✓	11	(5.03, -2.30)	8
	✓	30	(4.92, -2.608)	8
	✓	∞	(6.70, -2.45)	15
	-	3	(-1.77, -0.73)	15
	-	5	(5.01, -2.30)	7
	-	11	(4.98, -2.24)	6
	-	30	(4.61, -2.55)	15
	-	∞	(8.09, -3.63)	15

Table 3: Robustness test results: the simulated shift was $(\Delta x, \Delta y) = (5.0, -2.33)$. Accurate registration is achieved under severe noisy conditions by using the projection operator.

5 Relationship to Gradient Methods

The proposed algorithm possesses high sub-pixel registration accuracy similar to the GM algorithm, while exhibiting much superior convergence range: GM methods can estimate translations up to 5 pixels, while the UPC can estimate translations up to 50% of the common support as demonstrated in Table 1.

In order to analytically compare the proposed algorithms to GM algorithm, a frequency analysis of the GM is given in section 7 leading to Eq. 7.6, whose solution is valid under the assumption of Eq. 7.6. Comparing equations 3.6 and 7.6:

$$\begin{aligned} \text{GM: } \quad \omega_x \Delta x + \omega_y \Delta y &= \exp \left(i \left(\omega_x \widetilde{\Delta x} + \omega_y \widetilde{\Delta y} \right) \right) - 1 \approx \omega_x \widetilde{\Delta x} + \omega_y \widetilde{\Delta y} \\ \text{UPC: } \quad \omega_x \Delta x + \omega_y \Delta y &= \varphi(i, j) = \omega_x \widetilde{\Delta x} + \omega_y \widetilde{\Delta y} \end{aligned} \tag{5.1}$$

where $(\widetilde{\Delta x}, \widetilde{\Delta y})$ are the measurements and $(\Delta x, \Delta y)$ are the estimated translation values.

We come to the conclusion that both algorithms solve the same equation, but while the GM algorithm solves it under an assumption that we have small translation using a Taylor series approximation, the UPC algorithm solves it directly without using any approximations, yielding the extended convergence range.

6 Summary and Conclusions

In this paper we proposed a new approach to robust sub-pixel image registration. The algorithm extends the Phase Correlation method to sub-pixel accuracy and was proven to have superior convergence range and accuracy of $O(10^{-3})$ compared to prior works [2,

6]. By calculating the correlation function's phase, we derived exact expressions for least-square estimation of the phase-shift without resulting to approximations, thus making the proposed algorithm robust to a wide range of illumination changes, blurring and noise. The least-square problem was solved by either the regular or the robust least-square algorithms depending on the level of present noise. The experimental results were backed up by analytic analysis establishing a relation to prior works.

7 Appendix A: Gradient method frequency analysis

Following the formulation used in [6] we analyze the Brightness Change Constraint Equation in the frequency domain:

$$I^2 - I^1 = I_x^1 \Delta x + I_y^1 \Delta y \quad (7.1)$$

where I^1 and I^2 are two images having some common support and I_x^1 , I_y^1 are the spatial derivative of I^1 .

$$\widehat{I}^2 - \widehat{I}^1 = \widehat{I}^1 i \omega_x \Delta x + \widehat{I}^1 i \omega_y \Delta y = \widehat{I}^1 i (\omega_x \Delta x + \omega_y \Delta y) \quad (7.2)$$

$$\frac{\widehat{I}^2}{\widehat{I}^1} - 1 = i (\omega_x \Delta x + \omega_y \Delta y) \quad (7.3)$$

Assuming translational motion between I^2 and I^1 and using Eq. 2.5 we have:

$$\exp\left(i\left(\omega_x\widetilde{\Delta x} + \omega_y\widetilde{\Delta y}\right)\right) - 1 = i\left(\omega_x\Delta x + \omega_y\Delta y\right) \quad (7.4)$$

where $(\widetilde{\Delta x}, \widetilde{\Delta y})$ are the measurements and $(\Delta x, \Delta y)$ are the estimated values. By taking the first order Taylor expansion of the left term in Eq. 7.4 around $(0, 0)$ we get:

$$\exp(x) = 1 + x + x^2 + \dots x^n \approx 1 + x \quad (7.5)$$

$$\omega_x\widetilde{\Delta x} + \omega_y\widetilde{\Delta y} = \omega_x\Delta x + \omega_y\Delta y \quad (7.6)$$

8 Appendix B: Software Availability

The latest GNU Public Licensed (GPL) version of the software is available for download at:

<http://www.math.tau.ac.il/~keller>

References

- [1] B. Porat, **A Course in Digital Signal Processing**, John Wiley Pub.,1997.
- [2] H. Shekarforoush, M. Berthod, and J. Zerubia, "Subpixel image registration by estimating the polyphase decomposition of cross power spectrum", in Proceedings 1996 IEEE

Computer Society Conference on Computer Vision and Pattern Recognition, USA, June 1996, pp. 532–537.

- [3] DuMouchel, W.H., and F.L. O’Brien (1989), “Integrating a robust option into a multiple regression computing environment”, Computer Science and Statistics: Proceedings of the 21st Symposium on the Interface, Alexandria, VA: American Statistical Association.
- [4] A. Averbuch, R. Coifman, D. Donoho, M. Israeli, Y. Keller “Polar FFT, RectoPolar FFT, an Applications”, in preperation.
- [5] Robert Frischholz, “Class of algorithms for realtime sub-pixel registration”
- [6] L. Barron, D. J. Fleet and S. S. Beauchemin, “Performance of Optical flow Techniques”, Int. J. Computational Vision, Vol. 12, pp. 43-77, 1994.
- [7] C. D. Kuglin and D. C. Hines. “The phase correlation image alignment method” In IEEE 1975 Conference on Cybernetics and Society, pages 163–165, September 1975.
- [8] M. Unser, P. Thévenaz, L.P. Yaroslavsky, “Convolution-Based Interpolation for Fast, High-Quality Rotation of Images”, IEEE Transactions on Image Processing, vol. 4, no. 10, pp. 1371-1381, October 1995.